DOCUMENT RESUME

ED 410 866 HE 030 462

AUTHOR Antons, Christopher M.; Dilla, Benjamin L.; Fultz, Miriam L.
TITLE Assessing Student Attitudes: Computer versus Pencil. AIR 1997 Annual Forum Paper.
PUB DATE 1997-05-00
NOTE 19p.; Paper presented at the Annual Forum of the Association for Institutional Research (37th, Orlando, FL, May 18-21, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attitude Measures; College Environment; *College Students; Comparative Analysis; *Computer Assisted Testing; Correlation; Evaluation Methods; Higher Education; Military Schools; Questionnaires; Racial Attitudes; Religion; Sex Bias; *Student Attitudes; *Student Evaluation; Test Validity; Undergraduate Students
IDENTIFIERS Air Force Academy CO; *AIR Forum

ABSTRACT

This study compared results of measuring student attitudes regarding race, gender, and religious climate using pencil-and-paper administration in a group setting with computer-administrated measurement over a local area network. Approximately 2,000 U.S. Air Force Academy cadets in February 1996 and 1,400 in November 1996 were administered the Cadet Social Climate Survey. The survey instrument contained demographic questions, Likert-scaled items regarding attitudes about school climate issues, and questions about personal experiences. The instrument used in November was modified to include items from the Balanced Inventory of Desirable Responding. Computer administration took less time for a student to complete, could be done in the privacy of the student's room, was easier to administer, and provided a ready data set for analysis. The comparison showed a significantly higher rate of return of valid responses using the computer version and found no significant differences in the pattern of responses between the two versions. However, significant differences were found on socially desirable responding in that computer administration appeared to encourage a socially desirable response set, probably due to concerns regarding anonymity. (Contains 14 references.) (JLS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Running Head: ASSESSING STUDENT ATTITUDES: COMPUTER VERSUS PENCIL

ED 410 866

Assessing Student Attitudes: Computer Versus Pencil-and-Paper Administration

Christopher M. Antons

Personnel Research Psychologist

Institutional Research and Assessment

United States Air Force Academy

HQ USAFA/XPR

2304 Cadet Drive, Suite 300

USAF Academy, CO 80840

(719) 333-3092

Benjamin L. Dilla

Director, Institutional Research

Westark Community College

Vines 210 P.O. Box 3649

Fort Smith, AR 72913-3649

(501) 788-7038

Miriam L. Fultz

Survey Research Administrator

Institutional Research and Assessment

United States Air Force Academy

HQ USAFA/XPR

2304 Cadet Drive, Suite 300

USAF Academy, CO, 80840

(719) 333-2587

IR 030 462

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

AIR

for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Seventh Annual Forum of the Association for Institutional Research held in Orlando, Florida, May 18-21, 1997. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Jean Endo
Editor
AIR Forum Publications

BEST COPY AVAILABLE

Abstract

Student attitudes on racial, gender and religious climate were compared using pencil-and-paper administration in a group setting versus computer administration over a local area network to on-campus dormitory rooms. Computer administration takes less time for the student to complete, can be done in the privacy of their room, is easier to administer and provides a ready data set for analysis. Significantly higher return of valid responses on the computer version and no significant difference in the pattern of responses between the two versions were found. Additionally, effects of the survey form on socially desirable responding were tested and significant differences were found.

39

Assessing Student Attitudes: Computer Versus Pencil-and-Paper Administration

Colleges and universities survey their students for a variety of reasons – to assess and improve education, to measure attitudes and values, to gather demographic and opinion data, and to obtain feedback on student services. As in the realm of student testing, computers seem to offer the means to make surveys easier and more efficient to conduct. Some of the obvious advantages include: saving paper (and the handling thereof), ease of data collection, and creation of an electronic data file allowing analysis efforts to begin almost immediately.

From a student perspective, the advantages may seem less obvious, although we suspect the computer form is somewhat more engaging and takes less time to complete. Access to computers may be a constraint, depending on the college or university. But, of more importance to the student, is fear of losing the anonymity of their responses and trust in any assurances of anonymity.

The more sensitive the issue being surveyed, the more important anonymity of the survey respondent is to the collection of valid data. On paper surveys, students may feel their input is anonymous as long as name or other unique identifiers are not requested and sample sizes are relatively large. However, paper surveys at our institution are frequently administered in an auditorium or group setting. Anecdotal evidence from past respondents indicates there is some perceived pressure to complete the surveys and some fear of reprisal when responses are given in a group setting. Thus, because all students attending this institution have a computer in their room and all students are housed in residence halls on campus, one advantage of administering a survey on a local computer network is enhanced privacy for the individual respondent. In our early attempts at network administration,

feedback from the students indicated that at least some students felt more secure responding to the sensitive topic of student social climate in this more private setting.

Most studies of computer testing are concerned with the ability to quickly home in on aptitude or achievement level (e.g., Van de Vijver & Harsveld, 1984) or are used as a decision-making tool. In these studies, the emphasis is on metric form equivalence. Van Sickle, Kimmel and Kapes. (1993) established equivalence of forms between computer based and paper versions of the Strong-Campbell Interest Inventory and better test-retest reliability on the computer form, using American Psychological Association (APA, 1986) Guidelines for Computer-Based Tests. Comparisons of computer-based with paper administration of attitude inventories are sparse. Empirically, the results of these studies are mixed.

Using the Jenkins Activity Survey, Holden and Hickman (1987) found paper and computerized versions of the clinical instrument to be comparable with respect to means, variances, reliability and validity across the two forms. Their study was conducted with college students who were familiar and were competent with the use of computers. They suggested that equivalency must be demonstrated for different instruments and/or different populations.

Davis and Cowles (1989) compared responses to a battery of psychological tests on computer and paper forms in the context of social desirability. They found test-retest reliability better on the computer forms. However, they also found a "faking good" effect for the computer-based method. Those high on the Marlowe-Crown Social Desirability scale in the computer administered test group reported lower (better) scores on the neuroticism and anxiety scales.

The Marlowe-Crowne scale and its modifications use a single dimension of social desirable responding. Paulhus (1984) made a case for measuring two distinct dimensions. In a confirmatory factor analytic study, he found that the two dimensions of Impression Management (IM) and Self-Deception (SD) provided the best fit to the data. The first dimension measures the more social aspects of providing favorable responses in a more public or face-to-face situation and the second is a more personal enduring characteristic less likely to be influenced by the situation. He suggested IM be controlled for in self-reported personality inventories.

In a job applicant situation, Martin and Nagao (1989) found less socially desirable responding occurred with computer interviews than face-to-face interviews. Participants felt the computer was less obtrusive and their responses more private, since anonymity was not possible in a job interview. They suggested that people might respond more freely to a nonjudgmental machine.

Contrary to expectations, Latenschlager and Flaherty (1990) found a greater tendency toward a socially desirable response set with computer administration of two social desirability scales, when compared to traditional paper-and-pencil administration. Students in the computer administration mode were surveyed individually on a computer in a laboratory setting, while the paper mode included both individual and group settings. The individualized administration seemed to detract from the potentially greater privacy offered by computer administration. Of those surveyed individually (computer or paper, anonymous or identified, on the survey), a post-experiment questionnaire revealed greater concern that the experimenter might evaluate their responses after survey completion.

Booth-Kewley, Edwards and Rosenfeld (1992) conducted attitudinal surveys with Navy recruits in group settings using paper versus computer administration. For impression management (one aspect of socially desirable responding), they found no difference for mode of administration but did find a significant difference in the expected direction for identified versus anonymous respondents. Respondents in the anonymous condition, as well as those in the computer administration mode (regardless of condition), were significantly less likely to think it would be easy for someone to determine their individual responses.

The focus of this research effort is first, to establish the equivalence of computer and paper forms of a social attitude survey and second, to assess the degree to which student attitudes on the sensitive topics of racial, gender and religious climate are influenced by a social desirability response set in a computer form versus a paper form. The purpose of this investigation is to assess the effects of computer administration on responses to attitudinal questionnaires. Student attitudes on racial, gender and religious climate have been collected for a number of years at this institution. Recently, we administered a computer based survey on a local network to facilitate ease of administration and data collection. We were concerned both with its effect on valid responses and on social desirability response set. The expectation was that students would be more open to expressing their social attitudes on the computer than in a group setting with paper forms, thus lessening socially desirable responding. Therefore, the two forms of the survey should not only be statistically equivalent, but also less likely to solicit socially desirable responding because of the more private and anonymous conditions of the computer form.

Method

Research Participants

Approximately 2000 U.S. Air Force Academy cadets in February, 1996 and 1400 in November, 1996 were administered the Cadet Social Climate Survey. Students were randomly selected by squadron. A squadron was approximately 100 students housed together on the same dormitory floor. Previous analysis at our institution has shown that students are randomly assigned to squadron so that no differences in aptitude measures or demographics exists among squadrons. Therefore, the squadron unit was used as a cluster for random selection.

Materials

The survey instrument contained six demographic questions, 66 Likert-scaled questions to obtain attitudes about climate issues and 18 questions about personal experiences. The Likert-scaled questions used a 6-point scale from 1 = strongly disagree to 6 = strongly agree, with no neutral point. Some of the attitude questions were used as validity checks. The questions required an obviously false response. When students randomly responded to the paper version or pressed the default key on the computer version, they gave an invalid response to those items.

The survey instrument for the November, 1996 administration was modified to include 15 items taken from the Balanced Inventory of Desirable Responding, Version 6 (BIDR) (Paulhus, 1991). Both the Marlowe-Crowne Social Desirability Scale (1960) and BIDR inventories we judged too lengthy and likely to have hampered return rates by adding several minutes to response time. Hays, Hayashi, Toshi and Stewart (1989) developed a 5-

item scale taken primarily from the 33-item Marlowe-Crowne inventory. However, we felt the two dimensions of the BIDR were important to measure separately. Therefore, we selected 15 items most central to the two dimensions of the BIDR (Paulhus, 1991) and most topical to our military student population. Based on previous use of this inventory with our student population (Fultz, 1997), items with the strongest item-scale correlations were selected. Seven items were selected to measure Self-Deceptive Enhancement (BIDR-SDE). This scale measures the extent of inflated self-descriptions and exaggerated self-confidence. A sample item is "I never regret my decisions." Eight items were selected to measure Impression Management (BIDR-IM). This scale measures the extent of positive self-presentation, enhanced self-image in front of others and self, and the presentation of oneself as socially conventional. A sample item is "I don't gossip about other people's business." Both sets of selected items were balanced with positively and negatively scored items. Negatively scored items were reversed so that a higher score means greater levels of impression management and self-deception. The BIDR uses a 7-point scale Likert-type scale from 1 = not true to 7 = very true. The midpoint value, four, is anchored with somewhat true. In order to compare results to previous studies, scores were linearly converted to the same scale as each of the two full 20-item, continuously scored, BIDR scales (scale raw score range 20-140).

Procedure

The controlled (predictive) variable of interest was mode of administration. In February of 1996, the social climate survey was administered to about 1200 students on personal computers (PC's) in their dormitory rooms over a local area network (LAN).

Another 800 students simultaneously completed the same survey in a paper-and-pencil form in a group (auditorium) setting. Selection of the subjects and assignment to mode of administrations was random by cluster. The cluster was defined as a group of about 100 cadets organized in a squadron.

The climate survey was administered again in November 1996. As with the previous administration, the survey was administered to about 700 students individually on their PC's over the LAN. Another 700 students simultaneously completed the same survey in a paper-and-pencil form in a group (auditorium) setting. This survey included the additional BIDR items to measure socially desirable responding. Electronic files and scan sheets were collected and processed.

Instructions for assignment to mode of administration were given via electronic mail (e-mail). E-mail instructions to those taking the computer version included details on how to open the file, begin the survey and save the file. While students were housed two or three to a room and their personal computers were located on individual desks, they work separately and not side-by-side even within a room. Those assigned to the paper administration were instructed to go to an auditorium at the designated time.

Verbal instructions to the paper administration and written instructions on the first screen of the computer version explained the content of the survey, how to complete it, and that participation was voluntary. Included in the instructions was a statement about the anonymity of the individual responses. The Office of Institutional Research had exclusive access to the data and any reported results were with respect to groups and not individuals. No unique student identifiers were obtained on the data file and no combinations of

demographic information to break respondents down into small groups were used in any reports.

Results

For the February administration, 966 (81%) of the computer administered and 595 (74%) of the paper surveys were completed and returned. Both types of administrations were subjected to appropriate data screening. Comparing both versions, rates of completed surveys and valid response rates were significantly lower for the paper administration; $z(1995)=3.74$, $p<.05$. For this and all subsequent hypothesis testing, we used a type I (alpha) level of .05. The difference in valid response rates was also significant; $z(1559)=3.47$, $p<.05$. Twenty-six percent of completed paper surveys were eliminated due to invalid responses while only 17 percent of the computer version were invalid.

We also compared attitudinal responses and the factor structure of attitudinal items by mode of administration. Table 1 shows that, of the valid responses, a tests of differences between means revealed that attitudes on the dimensions of race ($t(747)=.61, p>.05$), gender ($t(1151)=.24, p>.05$), and religious ($t(1160)=.28, p>.05$) climate, were not significantly different by mode of administration. That is, we found no evidence that mode of administration resulted in biased self-reported climate attitudes. Further, variances were also not significantly different for gender ($F(1,1152)=.04, p<.05$) and religious($F(1,1161)=.04, p<.05$) climates, but was significantly different for the racial climate scale ($F(1,1154)=4.96, p>.05$). Psychometric properties, as measured by factor structure in principal components analysis (PCA) were nearly identical. Reliability indices (Cronbach's alpha) for each scale, were not significantly different for the two modes of administration

(Fs(1,1154, 1160 and 1151 respectively)=1.38, 1.07 and 1.15 respectively,p>.05) (Feldt, 1969).

Table 1

Properties of Attitude Scales by Mode of Administration: Feb 96 Survey

Scales	Mode	n	Difference in M	SD	Alpha
Racial	Paper	403	.021	.583	.792
	Computer	753		.521	.713
Gender	Paper	403	.012	.851	.801
	Computer	751		.831	.814
Religious	Paper	406	.014	.816	.689
	Computer	757		.780	.641

Results from the November, 1996 administration were quite different. Mode of administration did not affect the valid response rate, but did affect both average responses to self-reported attitudes and the measures of socially desirable responding. For the November, 1996 administration, 564 (81%) of the computer administered and 582 (83%) of the paper surveys were completed and returned. Those proportions were not significantly different z(1395)=0.95, p>05. Valid responses were much higher this time by either mode. Nine percent of completed paper surveys and 8% of computer version were eliminated due to invalid responses, resulting in 76% and 74% valid return rates, respectively. Returns and valid returns did not differ significantly by mode of administration, z(1144)=1.16, p>.05 .

Mean differences on Table 2 indicate those on the computer version espoused more positive climate attitudes. Across gender (t(984)=2.79,p<.05), race (t(882)=3.00,p<.05), and religious (t(1000)=2.42,p<.05) climate attitudes, those responding to the computer version expressed significantly more favorable attitudes, or at least a stronger inclination to positively inflate their responses.

Table 2

Properties of Attitude Scales by Mode of Administration: Nov 96 Survey

Scales	Mode	n	Difference in M	SD	Alpha
Racial	Paper	500	.123	.756	.830
	Computer	488		.516	.730
Gender	Paper	501	.147	.870	.778
	Computer	490		.789	.783
Religious	Paper	510	.127	.858	.640
	Computer	494		.802	.700

As in the February, 1996 administration, results of factor analysis and reliability indices were similar for both forms of the survey. Reliability indices (Cronbach's alpha) for each scale, were not significantly different for the two modes of administration (Fs(1,1002, 986 and 989 respectively)=1.02, 1.59 and 1.20 respectively,p>.05) (Feldt, 1969). Therefore, while psychometric measures exhibited acceptable levels of validity and reliability of climate dimensions for both computer and paper forms, those taking the computer form of the survey expressed more favorable climate attitudes.

However, those taking the computer version gave more inflated self-descriptions of both Self-Deceptive Enhancement and Impression Management as measured by the abbreviated sub-scales of the BIDR (all scores were converted to the full 20-140 scale for both BIDR sub-scales). Table 3 shows the mean scores for each scale for each mode of administration. Mean scores were significantly higher on both scales by computer mode of administration (Impression Management, t(976)=7.50, p<.05, and Self-Deception, t(971)=4.77, p<.05).

Table 3

Properties of BIDR Scales by Mode of Administration

Scales	Mode	n	M	SD	Alpha
Impression Management	Paper	495	87.34	17.92	.631
	Computer	497	96.45	20.27	.682
Self Deception	Paper	479	91.32	17.51	.607
	Computer	494	96.85	18.66	.607

Discussion

Statistically, we were satisfied that the two forms of our survey were equivalent. The dimensions of interest were similar under both forms and equally ~~equally~~ internally consistent, as measured by the reliability coefficients. However, we found differences in tendencies to report socially desirable responses.

Results support evidence in the literature that computer administration of an attitude assessment, particularly if their is concern for anonymity of responses, encourages a socially desirable response set. Our results are similar to Lautenschlager, et. al. (1991) in that computer administration resulted in higher, not lower, socially desirable responding. However, they found a larger difference for the Self-Deception scale than Impression Management. We found the opposite. The larger difference in Impression Management is more in keeping with findings from the original development of the BIDR (Paulhus, 1984). Impression Management is more likely to be influenced by social conditions and lack of anonymity is clearly one of those conditions.

Anonymity became a more crucial factor influencing response set on the climate surveys. In the February 1996 administration, students reported feeling more anonymous completing the surveys in the privacy of their rooms. Past studies have shown that

anonymity significantly affects the social desirability response set, particularly the Impression Management scale (Booth-Kewley, et. al, 1992). Expectations of being identified increase the likelihood of socially-desirable responding.

In the November, 1996 administration, timing of events created a loss in credibility concerning anonymity. Recent concerns on campus about appropriate use of the local area network by students and new network security actions may have made students feel their survey responses on the computer network might be scrutinized. This untimely event may alone have been responsible for differences in responses by mode of administration. While the software used does not include individual identifiers with responses, students do log onto the network and into the program using unique account names and passwords. Despite written comments from students to the contrary, a measure of socially desirable responding indicated a significant effect for mode of administration. This difference must either be accounted for in estimating student attitudes or paper-and-pencil should be the preferred mode of administration.

Administratively, computer surveys have many advantages. The paper-less environment and electronic data collection of multiple choice responses as well as written comments saves countless hours in processing time. With pre-established reporting procedures, survey results can be rapidly returned to students and administrators. Fast feedback improves student interest in completing the surveys. The concern, particularly on attitude surveys, is that method of administration will influence reporting. In particular, socially desirable responses may be encouraged under some conditions. From comments on the February, 1996 administration, despite concerns for anonymity in a computer

environment, students felt more privacy when using the computer in their rooms. Untimely events on the November, 1996 administration, when concerns for anonymity may have overwhelmed the benefits of privacy, appeared to cancel that positive effect.

It is hoped that, as well as easing the administrative burden, computer collection of attitude surveys will actually lessen the probability of socially desirable responding by student participants. This can only come about if students are confident their responses are anonymous. Such confidence is developed over time and is more easily lost than gained. Future developments in survey software and network technologies may increase students' feelings of security that their individual responses to a computerized survey are both private and anonymous. We continue to track our survey efforts in both paper and computer form until we can establish the ideal conditions under which the computer form provides us with not only efficiently collected data, but also valid and reliable data.

References

American Psychological Association. (1986). Guidelines for Computer-Based Tests and Interpretations. Washington, DC: Author.

Booth-Kewley, S., Edwards, J. E. & Rosenfeld, P. (1992). Impression management, social desirability, and computer administration of attitude questionnaires: does the computer make a difference? Journal of Applied Psychology, 77, 562-566.

Davis, C. & Cowles, M. (1989). Automated psychological testing: method of administration, need for approval, and measures of anxiety. Educational and Psychological Measurement, 49, 311-320.

Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. Psychometrika, 34, 363-373.

Fultz, M. L. (1997). Predicting voluntary turnover: the efficacy of between- versus within-subject predictions in an application of the theory of planned behavior. Unpublished doctoral dissertation, California School of Professional Psychology, San Diego.

Hays, R. D. , Hayashi, T. & Stewart, A. L. (1989). A five-item measure of socially desirable response set. Educational and Psychological Measurement, 49, 629-636.

Holden, R. R. & Hickman, D. (1987). Computerized versus standard administration of the Jenkins activity survey (form T). Journal of Human Stress, Winter 1987, 175-179.

Lautenschlager, G. J. & Flaherty, V. L. (1990). Computer administration of questions: more desirable or more social desirability? Journal of Applied Psychology, 75, 310-314.

Martin, C.L. & Nagao, D.H. (1989). Some effects of mode of test administration on test performance. Educational and Psychological Measurement, 46, 467-474.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. Journal of Personality and Social Psychology, 46, 598-609.

Paulhus, D. L. (1991). BIDR Reference Manual For Version 6. Unpublished manuscript, University of British Columbia.

Paulhus, D. L. & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. Journal of Personality and Social Psychology, 60, 307-317.

Van de Vijver, F. J. R. & Harsveld, M. (1994). The incomplete equivalence of the paper-and-pencil and computerized versions of the general aptitude test battery. Journal of Applied Psychology, 79, 852-859.

Vansickle, T. R., Kimmel, C. & Kapes, J. T. (1993). Test-retest equivalency of the computer-based and paper-pencil versions of the Strong-Campbell interest inventory. Computers in Human Behavior, 9, 401-409.